(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
24 January 2002 (24.01.2002)

PCT

(10) International Publication Number
WO 02/07151 A2

(51) International Patent Classification[7]: G10L 21/00

(21) International Application Number: PCT/US01/22490

(22) International Filing Date: 17 July 2001 (17.07.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/219,297 19 July 2000 (19.07.2000) US
09/905,361 12 July 2001 (12.07.2001) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:
US 60/219,297 (CIP)
Filed on 19 July 2000 (19.07.2000)
US 09/905,361 (CIP)
Filed on 12 July 2001 (12.07.2001)

(71) Applicant *(for all designated States except US)*: ALIPHCOM [US/US]; 2nd floor, 54 Mint Street, San Francisco, CA 94103 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: BURNETT, Gregory, C. [US/US]; 637 South H Street, Livermore, CA 94550 (US). BREITFELLER, Eric, F. [US/US]; 11526 Streambed Place, Dublin, CA 94568-2294 (US).

(74) Agent: PERKINS COIE, LLP; P.O. Box 2168, Menlo Park, CA 94026 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/07151 A2

(54) Title: METHOD AND APPARATUS FOR REMOVING NOISE FROM ELECTRONIC SIGNALS

(57) Abstract: A method and system are provided for acoustic noise removal from human speech, wherein noise is removed without respect to noise type, amplitude, or orientation. The system includes microphones and a voice activity detection (VAD) data stream coupled among a processor. The microphones receive acoustic signals and the VAD produces a signal including a binary one when speech (voiced and unvoiced) is occurring and a binary zero in the absence of speech. The processor includes denoising algorithms that generate transfer functions. The transfer functions include a transfer function generated in response to a determination that voicing information is absent from the received acoustic signal during a specified time period. The transfer functions also include transfer functions generated in response to a determination that voicing information is present in the acoustic signal during a specified time period. At least one denoised acoustic data stream is generated using the transfer functions.

METHOD AND APPARATUS FOR REMOVING NOISE FROM ELECTRONIC SIGNALS

FIELD OF THE INVENTION

The invention is in the field of mathematical methods and electronic systems for removing or suppressing undesired acoustical noise from acoustic transmissions or recordings.

BACKGROUND

In a typical acoustic application, speech from a human user is recorded or stored and transmitted to a receiver in a different location. In the environment of the user, there may exist one or more noise sources that pollute the signal of interest (the user's speech) with unwanted acoustic noise. This makes it difficult or impossible for the receiver, whether human or machine, to understand the user's speech. This is especially problematic now with the proliferation of portable communication devices like cellular telephones and personal digital assistants. There are existing methods for suppressing these noise additions, but they either require far too much computing time or cumbersome hardware, distort the signal of interest too much, or lack in performance to be useful. Many of these methods are described in textbooks such as "Advanced Digital Signal Processing and Noise Reduction" by Vaseghi, ISBN 0-471-62692-9. Consequently, there is a need for noise removal and reduction methods that address the shortcomings of typical systems and offer new techniques for cleaning acoustic signals of interest without distortion.

SUMMARY

A method and system are provided for acoustic noise removal from human speech, wherein the noise can be removed and the signal restored without respect to noise type, amplitude, or orientation. The system includes microphones and sensors coupled with a processor. The microphones receive acoustic signals including both noise and speech signals from human signal sources. The sensors yield a binary Voice Activity Detection (VAD) signal that provides a signal that is a binary "1" when speech (both voiced and unvoiced) is occurring and a binary "0" when no speech is occurring. The VAD signal can be obtained in numerous ways, for example, using acoustic gain, accelerometers, and radio frequency (RF) sensors.

The processor system and method includes denoising algorithms that calculate the transfer function among the noise sources and the microphones as well as the transfer function among the human user and the microphones. The transfer functions are used to remove noise from the received acoustic signal to produce at least one denoised acoustic data stream.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a block diagram of a denoising system of an embodiment.

Figure 2 is a block diagram of a noise removal algorithm of an embodiment, assuming a single noise source and a direct path to the microphones.

Figure 3 is a block diagram of a front end of a noise removal algorithm of an embodiment, generalized to *n* distinct noise sources (these noise sources may be reflections or echoes of one another).

Figure 4 is a block diagram of a front end of a noise removal algorithm of an embodiment in the most general case where there are *n* distinct noise sources and signal reflections.

Figure 5 is a flow diagram of a denoising method of an embodiment.

Figure 6 shows results of a noise suppression algorithm of an embodiment for an American English female speaker in the presence of airport terminal noise that includes many other human speakers and public announcements.

DETAILED DESCRIPTION

Figure 1 is a block diagram of a denoising system of an embodiment that uses knowledge of when speech is occurring derived from physiological information on voicing activity. The system includes microphones 10 and sensors 20 that provide signals to at least one processor 30. The processor includes a denoising subsystem or algorithm.

Figure 2 is a block diagram of a noise removal system/algorithm of an embodiment, assuming a single noise source and a direct path to the microphones. The noise removal system diagram includes a graphic description of the process of an embodiment, with a single signal source (100) and a single noise source (101). This algorithm uses two microphones, a "signal" microphone (MIC 1, 102) and a "noise" microphone (MIC 2, 103), but is not so limited. MIC 1 is assumed to capture mostly signal with some noise, while MIC 2 captures mostly noise with some signal. This is the common configuration with conventional advanced acoustic systems. The data from the signal to MIC 1 is denoted by s(n), from the signal to MIC 2 by $s_2(n)$, from the noise to MIC 2 by n(n), and from the noise to MIC 1 by $n_2(n)$. Similarly, the data from MIC 1 is denoted by $m_1(n)$, and the data from MIC 2 $m_2(n)$, where s(n) denotes a discrete sample of the analog signal from the source.

The transfer functions from the signal to MIC 1 and from the noise to MIC 2 are assumed to be unity, but the transfer function from the signal to MIC 2 is denoted by $H_2(z)$ and from the noise to MIC 1 by $H_1(z)$. The assumption of unity transfer functions does not inhibit the generality of this algorithm, as the actual relations between the signal, noise, and microphones are simply ratios and the ratios are redefined in this manner for simplicity.

In conventional noise removal systems, the information from MIC 2 is used to attempt to remove noise from MIC 1. However, an unspoken assumption is that the Voice Activity Detection (VAD) is never perfect, and thus the denoising must be performed cautiously, so as not to remove too much of the signal along with the noise. However, if the VAD is assumed to be perfect and is equal to zero when there is no speech being produced by the user,

and one when speech is produced, a substantial improvement in the noise removal can be made.

In analyzing the single noise source and direct path to the microphones, with reference to **Figure 2,** the acoustic information coming into MIC 1 is denoted by $m_1(n)$. The information coming into MIC 2 is similarly labeled $m_2(n)$. In the *z* (digital frequency) domain, these are represented as $M_1(z)$ and $M_2(z)$. Then

$$M_1(z) = S(z) + N_2(z)$$
$$M_2(z) = N(z) + S_2(z)$$

with

$$N_2(z) = N(z)H_1(z)$$
$$S_2(z) = S(z)H_2(z)$$

so that

$$M_1(z) = S(z) + N(z)H_1(z)$$
$$M_2(z) = N(z) + S(z)H_2(z) \qquad \text{Eq. 1}$$

This is the general case for all two microphone systems. In a practical system there is always going to be some leakage of noise into MIC 1, and some leakage of signal into MIC 2. Equation 1 has four unknowns and only two known relationships and therefore cannot be solved explicitly.

However, there is another way to solve for some of the unknowns in Equation 1. The analysis starts with an examination of the case where the signal is not being generated, that is, where the VAD signal equals zero and speech is not being produced. In this case, s(n) = S(z) = 0, and Equation 1 reduces to

$$M_{1n}(z) = N(z)H_1(z)$$
$$M_{2n}(z) = N(z)$$

where the *n* subscript on the M variables indicate that only noise is being received. This leads to

$$M_{1n}(z) = M_{2n}(z)H_1(z)$$
$$H_1(z) = \frac{M_{1n}(z)}{M_{2n}(z)} \qquad \text{Eq. 2}$$

$H_1(z)$ can be calculated using any of the available system identification algorithms and the microphone outputs when the system is certain that only noise is being received. The calculation can be done adaptively, so that the system can react to changes in the noise.

A solution is now available for one of the unknowns in Equation 1. Another unknown, $H_2(z)$, can be determined by using the instances where the VAD equals one and speech is being produced. When this is occurring, but the recent (perhaps less than 1 second) history of the microphones indicate low levels of noise, it can be assumed that $n(s) = N(z) \sim 0$. Then Equation 1 reduces to

$$M_{1s}(z) = S(z)$$
$$M_{2s}(z) = S(z)H_2(z)$$

which in turn leads to

$$M_{2s}(z) = M_{1s}(z)H_2(z)$$
$$H_2(z) = \frac{M_{2s}(z)}{M_{1s}(z)}$$

which is the inverse of the $H_1(z)$ calculation. However, it is noted that different inputs are being used – now only the signal is occurring whereas before only the noise was occurring. While calculating $H_2(z)$, the values calculated for $H_1(z)$ are held constant and vice versa. Thus, it is assumed that $H_1(z)$ and $H_2(z)$ do not change substantially while the other is being calculated.

After calculating $H_1(z)$ and $H_2(z)$, they are used to remove the noise from the signal. If Equation 1 is rewritten as

$$S(z) = M_1(z) - N(z)H_1(z)$$
$$N(z) = M_2(z) - S(z)H_2(z)$$
$$S(z) = M_1(z) - [M_2(z) - S(z)H_2(z)]H_1(z) \text{ ,}$$
$$S(z)[1 - H_2(z)H_1(z)] = M_1(z) - M_2(z)H_1(z)$$

then N(z) may be substituted as shown to solve for S(z) as:

$$S(z) = \frac{M_1(z) - M_2(z)H_1(z)}{1 - H_2(z)H_1(z)} \qquad \text{Eq. 3}$$

If the transfer functions $H_1(z)$ and $H_2(z)$ can be described with sufficient accuracy, then the noise can be completely removed and the original signal recovered. This remains true without respect to the amplitude or spectral

characteristics of the noise. The only assumptions made are a perfect VAD, sufficiently accurate $H_1(z)$ and $H_2(z)$, and that $H_1(z)$ and $H_2(z)$ do not change substantially when the other is being calculated. In practice these assumptions have proven reasonable.

The noise removal algorithm described herein is easily generalized to include any number of noise sources. **Figure 3** is a block diagram of a front end of a noise removal algorithm of an embodiment, generalized to *n* distinct noise sources. These distinct noise sources may be reflections or echoes of one another, but are not so limited. There are several noise sources shown, each with a transfer function, or path, to each microphone. The previously named path H_2 has been relabeled as H_0, so that labeling noise source 2's path to MIC 1 is more convenient. The outputs of each microphone, when transformed to the z domain, are:

$$\begin{aligned} M_1(z) &= S(z) + N_1(z)H_1(z) + N_2(z)H_2(z) + \ldots N_n(z)H_n(z) \\ M_2(z) &= S(z)H_0(z) + N_1(z)G_1(z) + N_2(z)G_2(z) + \ldots N_n(z)G_n(z) \end{aligned} \qquad \text{Eq. 4}$$

When there is no signal (VAD = 0), then (suppressing the z's for clarity)

$$\begin{aligned} M_{1n} &= N_1H_1 + N_2H_2 + \ldots N_nH_n \\ M_{2n} &= N_1G_1 + N_2G_2 + \ldots N_nG_n \end{aligned} \qquad \text{Eq. 5}$$

A new transfer function can now be defined, analogous to $H_1(z)$ above:

$$\tilde{H}_1 = \frac{M_{1n}}{M_{2n}} = \frac{N_1H_1 + N_2H_2 + \ldots N_nH_n}{N_1G_1 + N_2G_2 + \ldots N_nG_n} \qquad \text{Eq. 6}$$

Thus $\tilde{H}_1$ depends only on the noise sources and their respective transfer functions and can be calculated any time there is no signal being transmitted. Once again, the *n* subscripts on the microphone inputs denote only that noise is being detected, while an *s* subscript denotes that only signal is being received by the microphones.

Examining Equation 4 while assuming that there is no noise produces

$$\begin{aligned} M_{1s} &= S \\ M_{2s} &= SH_0 \end{aligned}$$

Thus H_0 can be solved for as before, using any available transfer function calculating algorithm. Mathematically

$$H_0 = \frac{M_{2s}}{M_{1s}}$$

Rewriting Equation 4, using $\widetilde{H}_1$ defined in Equation 6, provides,

$$\widetilde{H}_1 = \frac{M_1 - S}{M_2 - SH_0} \qquad \text{Eq. 7}$$

Solving for S yields,

$$S = \frac{M_1 - M_2\widetilde{H}_1}{1 - H_0\widetilde{H}_1} \qquad \text{Eq. 8}$$

which is the same as Equation 3, with H_0 taking the place of H_2, and $\widetilde{H}_1$ taking the place of H_1. Thus the noise removal algorithm still is mathematically valid for any number of noise sources, including multiple echoes of noise sources. Again, if H_0 and $\widetilde{H}_1$ can be estimated to a high enough accuracy, and the above assumption of only one path from the signal to the microphones holds, the noise may be removed completely.

The most general case involves multiple noise sources and multiple signal sources. **Figure 4** is a block diagram of a front end of a noise removal algorithm of an embodiment in the most general case where there are *n* distinct noise sources and signal reflections. Here, reflections of the signal enter both microphones. This is the most general case, as reflections of the noise source into the microphones can be modeled accurately as simple additional noise sources. For clarity, the direct path from the signal to MIC 2 has changed from $H_0(z)$ to $H_{00}(z)$, and the reflected paths to Microphones 1 and 2 are denoted by $H_{01}(z)$ and $H_{02}(z)$, respectively.

The input into the microphones now becomes

$$\begin{aligned} M_1(z) &= S(z) + S(z)H_{01}(z) + N_1(z)H_1(z) + N_2(z)H_2(z) + \ldots N_n(z)H_n(z) \\ M_2(z) &= S(z)H_{00}(z) + S(z)H_{02}(z) + N_1(z)G_1(z) + N_2(z)G_2(z) + \ldots N_n(z)G_n(z) \end{aligned} \qquad \text{Eq. 9}$$

When the VAD = 0, the inputs become (suppressing the z's again)

$$\begin{aligned} M_{1n} &= N_1H_1 + N_2H_2 + \ldots N_nH_n \\ M_{2n} &= N_1G_1 + N_2G_2 + \ldots N_nG_n \end{aligned}$$

which is the same as Equation 5. Thus, the calculation of $\tilde{H}_1$ in Equation 6 is unchanged, as expected. In examining the situation where there is no noise, Equation 9 reduces to

$$M_{1s} = S + SH_{01}$$
$$M_{2s} = SH_{00} + SH_{02}$$

This leads to the definition of $\tilde{H}_2$:

$$\tilde{H}_2 = \frac{M_{2s}}{M_{1s}} = \frac{H_{00} + H_{02}}{1 + H_{01}} \qquad \text{Eq. 10}$$

Rewriting Equation 9 again using the definition for $\tilde{H}_1$ (as in Equation 7) provides

$$\tilde{H}_1 = \frac{M_1 - S(1 + H_{01})}{M_2 - S(H_{00} + H_{02})} \qquad \text{Eq. 11}$$

Some algebraic manipulation yields

$$S\left(1 + H_{01} - \tilde{H}_1(H_{00} + H_{02})\right) = M_1 - M_2\tilde{H}_1$$
$$S(1 + H_{01})\left[1 - \tilde{H}_1 \frac{(H_{00} + H_{02})}{(1 + H_{01})}\right] = M_1 - M_2\tilde{H}_1$$
$$S(1 + H_{01})\left[1 - \tilde{H}_1\tilde{H}_2\right] = M_1 - M_2\tilde{H}_1$$

and finally

$$S(1 + H_{01}) = \frac{M_1 - M_2\tilde{H}_1}{1 - \tilde{H}_1\tilde{H}_2} \qquad \text{Eq. 12}$$

Equation 12 is the same as equation 8, with the replacement of H_0 by $\tilde{H}_2$, and the addition of the $(1+H_{01})$ factor on the left side. This extra factor means that S cannot be solved for directly in this situation, but a solution can be generated for the signal plus the addition of all of its echoes. This is not such a bad situation, as there are many conventional methods for dealing with echo suppression, and even if the echoes are not suppressed, it is unlikely that they will affect the comprehensibility of the speech to any meaningful extent. The more complex calculation of $\tilde{H}_2$ is needed to account for the signal echoes in Microphone 2, which act as noise sources.

Figure 5 is a flow diagram of a denoising method of an embodiment. In operation, the acoustic signals are received 502. Further, physiological

information associated with human voicing activity is received 504. A first transfer function representative of the acoustic signal is calculated upon determining that voicing information is absent from the acoustic signal for at least one specified period of time 506. A second transfer function representative of the acoustic signal is calculated upon determining that voicing information is present in the acoustic signal for at least one specified period of time 508. Noise is removed from the acoustic signal using at least one combination of the first transfer function and the second transfer function, producing denoised acoustic data streams 510.

An algorithm for noise removal, or denoising algorithm, is described herein, from the simplest case of a single noise source with a direct path to multiple noise sources with reflections and echoes. The algorithm has been shown herein to be viable under any environmental conditions. The type and amount of noise are inconsequential if a good estimate has been made of $\tilde{H}_1$ and $\tilde{H}_2$, and if they do not change substantially while the other is calculated. If the user environment is such that echoes are present, they can be compensated for if coming from a noise source. If signal echoes are also present, they will affect the cleaned signal, but the effect should be negligible in most environments.

In operation, the algorithm of an embodiment has shown excellent results in dealing with a variety of noise types, amplitudes, and orientations. However, there are always approximations and adjustments that have to be made when moving from mathematical concepts to engineering applications. One assumption is made in Equation 3, where $H_2(z)$ is assumed small and therefore $H_2(z)H_1(z) \approx 0$, so that Equation 3 reduces to

$$S(z) \approx M_1(z) - M_2(z)H_1(z).$$

This means that only $H_1(z)$ has to be calculated, speeding up the process and reducing the number of computations required considerably. With the proper selection of microphones, this approximation is easily realized.

Another approximation involves the filter used in an embodiment. The actual $H_1(z)$ will undoubtedly have both poles and zeros, but for stability and

simplicity an all-zero Finite Impulse Response (FIR) filter is used. With enough taps (around 60) the approximation to the actual $H_1(z)$ is very good.

Regarding subband selection, the wider the range of frequencies over which a transfer function must be calculated, the more difficult it is to calculate it accurately. Therefore the acoustic data was divided into 16 subbands, with the lowest frequency at 50 Hz and the highest at 3700. The denoising algorithm was then applied to each subband in turn, and the 16 denoised data streams were recombined to yield the denoised acoustic data. This works very well, but any combinations of subbands (i.e. 4, 6, 8, 32, equally spaced, perceptually spaced, etc.) can be used and has been found to work as well.

The amplitude of the noise was constrained in an embodiment so that the microphones used did not saturate (i.e. operate outside a linear response region). It is important that the microphones operate linearly to ensure the best performance. Even with this restriction, very high signal-to-noise ratios (SNR) can be tested (down to about -10 dB).

The calculation of $H_1(z)$ was accomplished every 10 milliseconds using the Least-Mean Squares (LMS) method, a common adaptive transfer function. An explanation may be found in "Adaptive Signal Processing" (1985), by Widrow and Stearns, published by Prentice-Hall, ISBN 0-13-004029-0.

The VAD for an embodiment was derived from a radio frequency sensor and the two microphones, yielding very high accuracy (>99%) for both voiced and unvoiced speech. The VAD of an embodiment uses a radio frequency (RF) interferometer to detect tissue motion associated with human speech production, but is not so limited. It is therefore completely acoustic-noise free, and is able to function in any acoustic noise environment. A simple energy measurement can be used to determine if voiced speech is occurring. Unvoiced speech can be determined using conventional frequency-based methods, by proximity to voiced sections, or through a combination of the above. Since there is much less energy in unvoiced speech, its activation accuracy is not as critical as voiced speech.

With voiced and unvoiced speech detected reliably, the algorithm of an embodiment can be implemented. Once again, it is useful to repeat that the

noise removal algorithm does not depend on how the VAD is obtained, only that it is accurate, especially for voiced speech. If speech is not detected and training occurs on the speech, the subsequent denoised acoustic data can be distorted.

Data was collected in four channels, one for MIC 1, one for MIC 2, and two for the radio frequency sensor that detected the tissue motions associated with voiced speech. The data were sampled simultaneously at 40 kHz, then digitally filtered and decimated down to 8 kHz. The high sampling rate was used to reduce any aliasing that might result from the analog to digital process. A four-channel National Instruments A/D board was used along with Labview to capture and store the data. The data was then read into a C program and denoised 10 milliseconds at a time.

Figure 6 shows results of a noise suppression algorithm of an embodiment for an American English speaking female in the presence of airport terminal noise that includes many other human speakers and public announcements. The speaker is uttering the numbers 406-5562 in the midst of moderate airport terminal noise. The dirty acoustic data was denoised 10 milliseconds at a time, and before denoising the 10 milliseconds of data were prefiltered from 50 to 3700 Hz. A reduction in the noise of approximately 17 dB is evident. No post filtering was done on this sample; thus, all of the noise reduction realized is due to the algorithm of an embodiment. It is clear that the algorithm adjusts to the noise instantly, and is capable of removing the very difficult noise of other human speakers. Many different types of noise have all been tested with similar results, including street noise, helicopters, music, and sine waves, to name a few. Also, the orientation of the noise can be varied substantially without significantly changing the noise suppression performance. Finally, the distortion of the cleaned speech is very low, ensuring good performance for speech recognition engines and human receivers alike.

The noise removal algorithm of an embodiment has been shown to be viable under any environmental conditions. The type and amount of noise are inconsequential if a good estimate has been made of $\widetilde{H}_1$ and $\widetilde{H}_2$. If the user environment is such that echoes are present, they can be compensated for if

coming from a noise source. If signal echoes are also present, they will affect the cleaned signal, but the effect should be negligible in most environments.

Various embodiments are described herein with reference to the figures, but the detailed description and the figures are not intended to be limiting. Various combinations of the elements described have not been shown, but are within the scope of the invention which is defined by the following claims.

CLAIMS

What is claimed is:

1. A method for removing noise from acoustic signals, comprising:

receiving a plurality of acoustic signals;

receiving physiological information associated with human voicing activity;

generating at least one first transfer function representative of the plurality of acoustic signals upon determining that voicing information is absent from the plurality of acoustic signals for at least one specified period of time;

generating at least one second transfer function representative of the plurality of acoustic signals upon determining that voicing information is present in the plurality of acoustic signals for the at least one specified period of time;

removing noise from the plurality of acoustic signals using at least one combination of the at least one first transfer function and the at least one second transfer function to produce at least one denoised acoustic data stream.

2. The method of claim 1, wherein the plurality of acoustic signals include at least one reflection of at least one associated noise source signal and at least one reflection of at least one acoustic source signal.

3. The method of claim 1, wherein receiving physiological information comprises receiving physiological data associated with human voicing using at least one detector selected from a group consisting of radio frequency devices, electroglottographs, ultrasound devices, acoustic throat microphones, and airflow detectors.

4. The method of claim 1, wherein receiving the plurality of acoustic signals includes receiving using a plurality of independently located microphones.

5. The method of claim 1, wherein removing noise further includes generating at least one third transfer function using the at least one first transfer function and the at least one second transfer function.

6. The method of claim 1, wherein generating the at least one first transfer function comprises recalculating the at least one first transfer function during at least one prespecified interval.

7. The method of claim 1, wherein generating the at least one second transfer function comprises recalculating the at least one second transfer function during at least one prespecified interval.

8. The method of claim 1, wherein generating the at least one first transfer function and the at least one second transfer function comprises use of at least one technique selected from a group consisting of adaptive techniques and recursive techniques.

9. A method for removing noise from electronic signals, comprising:

detecting an absence of voiced information during at least one period;

receiving at least one noise source signal during the at least one period;

generating at least one transfer function representative of the at least one noise source signal;

receiving at least one composite signal comprising acoustic and noise signals; and

removing the noise signal from the at least one composite signal using the at least one transfer function to produce at least one denoised acoustic data stream.

10. The method of claim 9, wherein the at least one noise source signal includes at least one reflection of at least one associated noise source signal.

11. The method of claim 9, wherein the at least one composite signal includes at least one reflection of at least one associated composite signal.

12. The method of claim 9, wherein detecting comprises collecting physiological data associated with human voicing using at least one detector selected from a group consisting of radio frequency devices, electroglottographs, ultrasound devices, acoustic throat microphones, and airflow detectors.

13. The method of claim 9, wherein receiving includes receiving the at least one noise source signal using at least one microphone.

14. The method of claim 13, wherein the at least one microphone includes a plurality of independently located microphones.

15. The method of claim 9, wherein removing the noise signal from the at least one composite signal using the at least one transfer function includes generating at least one other transfer function using the at least one transfer function.

16. The method of claim 9, wherein generating at least one transfer function comprises recalculating the at least one transfer function during at least one prespecified interval.

17. The method of claim 9, wherein generating the at least one transfer function comprises calculating the at least one transfer function using at least one technique selected from a group consisting of adaptive techniques and recursive techniques.

18. A method for removing noise from electronic signals, comprising:

determining at least one unvoicing period during which voiced information is absent;

receiving at least one noise signal input during the at least one unvoicing period and generating at least one unvoicing transfer function representative of the at least one noise signal;

determining at least one voicing period during which voiced information is present;

receiving at least one acoustic signal input from at least one signal sensing device during the at least one voicing period and generating at least one voicing transfer function representative of the at least one acoustic signal;

receiving at least one composite signal comprising acoustic and noise signals; and

removing the noise signal from the at least one composite signal using at least one combination of the at least one unvoicing transfer function and the at least one voicing transfer function to produce at least one denoised acoustic data stream.

19. A system for removing noise from acoustic signals, comprising:

at least one receiver that receives at least one acoustic signal;

at least one sensor that receives physiological information associated with human voicing activity;

at least one processor coupled among the at least one receiver and the at least one sensor that generates a plurality of transfer functions, wherein at least one first transfer function representative of the at least one acoustic signal is generated in response to a determination that voicing information is absent from the at least one acoustic signal for at least one specified period of time, wherein at least one second transfer function representative of the at least one acoustic signal is generated in response to a determination that voicing information is present in the at least one acoustic signal for at least one specified period of time, wherein noise is removed from the at least one acoustic signal using at least one combination of the at least one first transfer function and the at least one second transfer function to produce at least one denoised acoustic data stream.

20. The system of claim 19, wherein the at least one sensor includes at least one radio frequency (RF) interferometer that detects tissue motion associated with human speech production.

21. The system of claim 19, wherein the at least one sensor includes at least one sensor selected from a group consisting of radio frequency devices, electroglottographs, ultrasound devices, acoustic throat microphones, and airflow detectors.

22. The system of claim 19, further comprising:

dividing acoustic data of the at least one acoustic signal into a plurality of subbands;

removing noise from each of the plurality of subbands using the at least one combination of the at least one first transfer function and the at least one second transfer function, wherein a plurality of denoised acoustic data streams are generated; and

combining the plurality of denoised acoustic data streams to generate the at least one denoised acoustic data stream.

23. The system of claim 19, wherein the at least one receiver includes a plurality of independently located microphones.

24. A system for removing noise from acoustic signals, comprising at least one processor coupled among at least one microphone and at least one voicing sensor, wherein the at least one voicing sensor collects physiological data associated with voicing, wherein an absence of voiced information is detected during at least one period using the at least one voicing sensor, wherein at least one noise source signal is received during the at least one period using the at least one microphone, wherein the at least one processor generates at least one transfer function representative of the at least one noise source signal, wherein the at least one microphone receives at least one composite signal comprising acoustic and noise signals, and the at least one processor removes the noise

signal from the at least one composite signal using the at least one transfer function to produce at least one denoised acoustic data stream.

25. A signal processing system coupled among at least one user and at least one electronic device, wherein the signal processing system includes at least one denoising subsystem for removing noise from acoustic signals, the denoising subsystem comprising at least one processor coupled among at least one receiver and at least one sensor, wherein the at least one receiver is coupled to receive at least one acoustic signal, wherein the at least one sensor is coupled to receive physiological information associated with human voicing activity, wherein the at least one processor generates a plurality of transfer functions, wherein at least one first transfer function representative of the at least one acoustic signal is generated in response to a determination that voicing information is absent from the at least one acoustic signal for at least one specified period of time, wherein at least one second transfer function representative of the at least one acoustic signal is generated in response to a determination that voicing information is present in the at least one acoustic signal for at least one specified period of time, wherein noise is removed from the at least one acoustic signal using at least one combination of the at least one first transfer function and the at least one second transfer function to produce at least one denoised acoustic data stream.

26. The signal processing system of claim 25, wherein the at least one electronic device includes at least one device selected from a group consisting of cellular telephones, personal digital assistants, portable communication devices, computers, video cameras, digital cameras, and telematics systems.

27. A computer readable medium comprising executable instructions which, when executed in a processing system, remove noise from received acoustic signals by:

receiving at least one acoustic signal;

receiving physiological information associated with human voicing activity;

generating at least one first transfer function representative of the at least one acoustic signal upon determining that voicing information is absent from the at least one acoustic signal for at least one specified period of time;

generating at least one second transfer function representative of the at least one acoustic signal upon determining that voicing information is present in the at least one acoustic signal for at least one specified period of time;

removing noise from the at least one acoustic signal using at least one combination of the at least one first transfer function and the at least one second transfer function to produce at least one denoised acoustic data stream.

28. An electromagnetic medium comprising executable instructions which, when executed in a processing system, remove noise from received acoustic signals by:

receiving at least one acoustic signal;

receiving physiological information associated with human voicing activity;

generating at least one first transfer function representative of the at least one acoustic signal upon determining that voicing information is absent from the at least one acoustic signal for at least one specified period of time;

generating at least one second transfer function representative of the at least one acoustic signal upon determining that voicing information is present in the at least one acoustic signal for at least one specified period of time;

removing noise from the at least one acoustic signal using at least one combination of the at least one first transfer function and the at least one second transfer function to produce at least one denoised acoustic data stream.

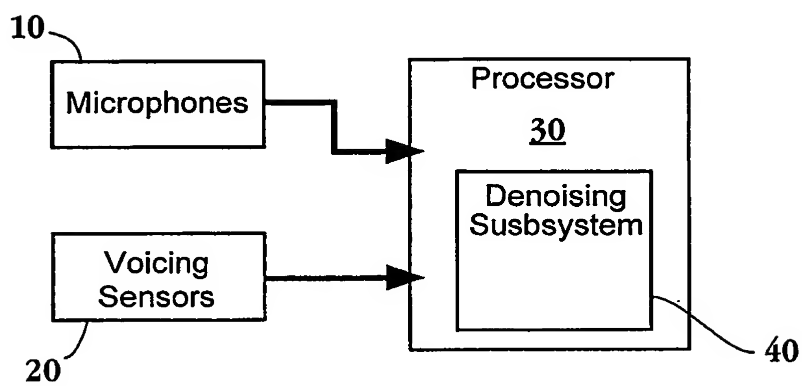
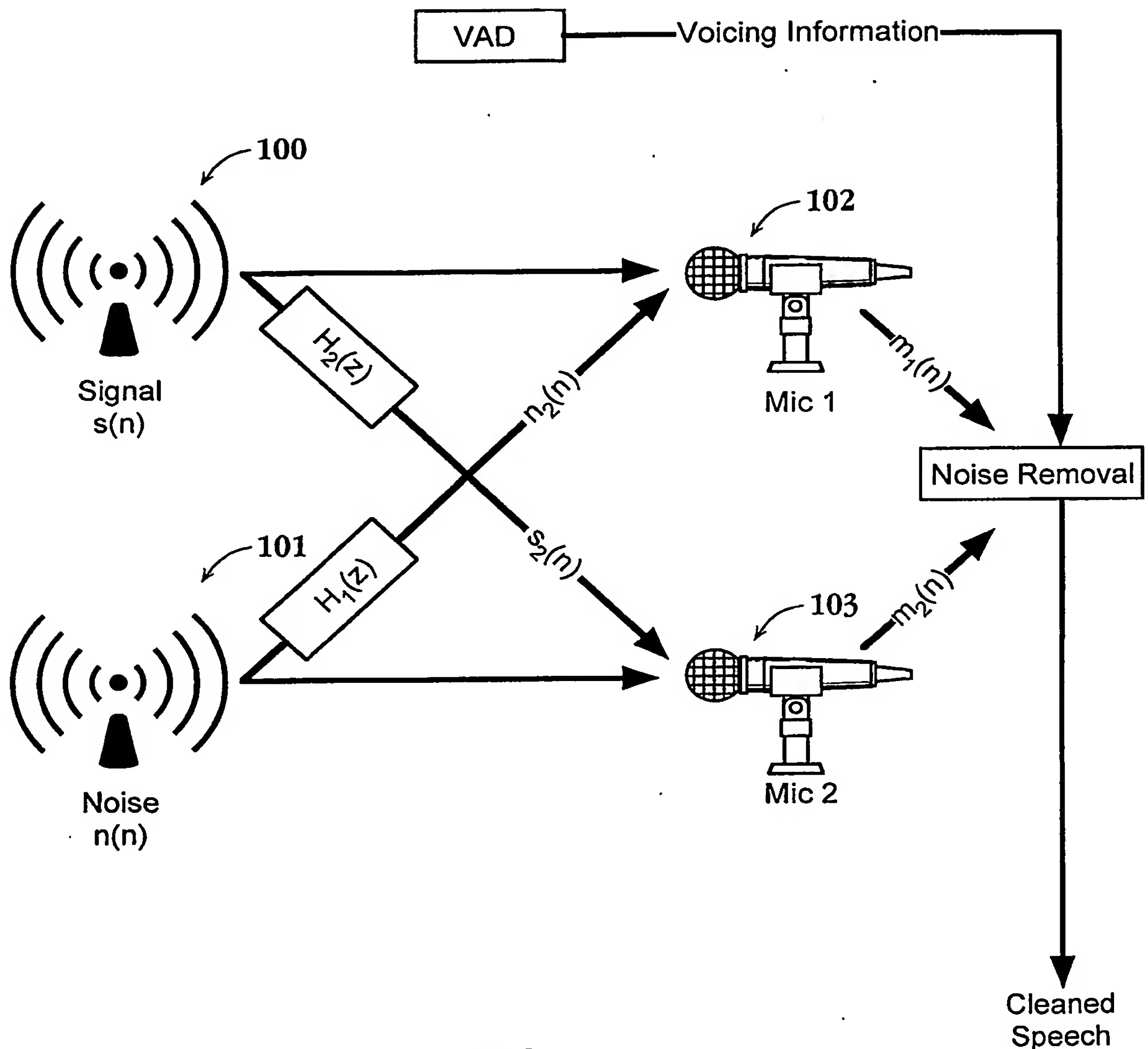

10
Microphones
Voicing
Sensors
20
Processor
30
Denoising
Susbsystem
40

Fig. 1



VAD
Voicing Information
100
Signal
s(n)
102
Mic 1
$H_2(z)$
$n_2(n)$
$m_1(n)$
Noise Removal
101
$H_1(z)$
$s_2(n)$
103
$m_2(n)$
Noise
n(n)
Mic 2
Cleaned
Speech

Fig. 2

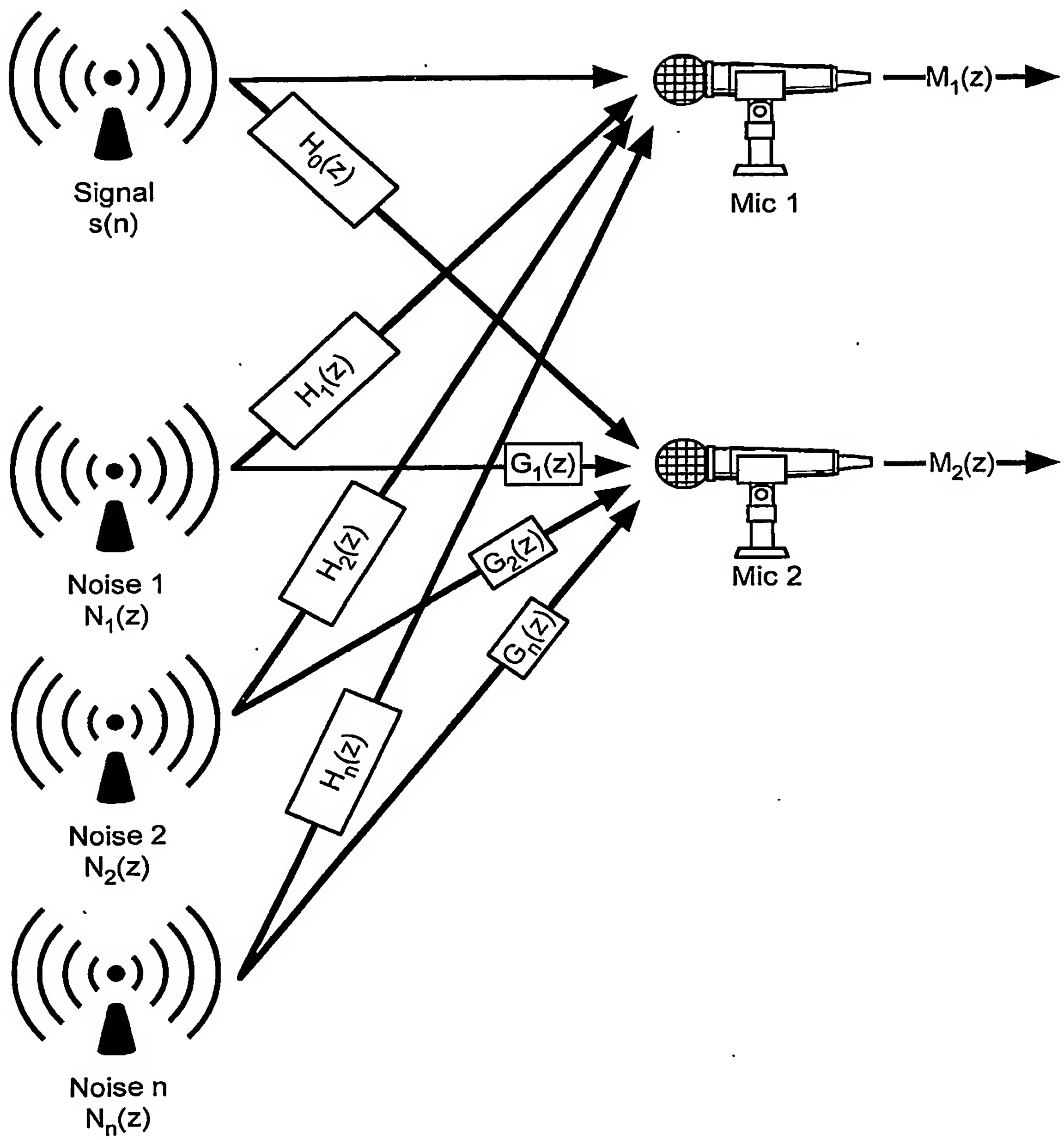
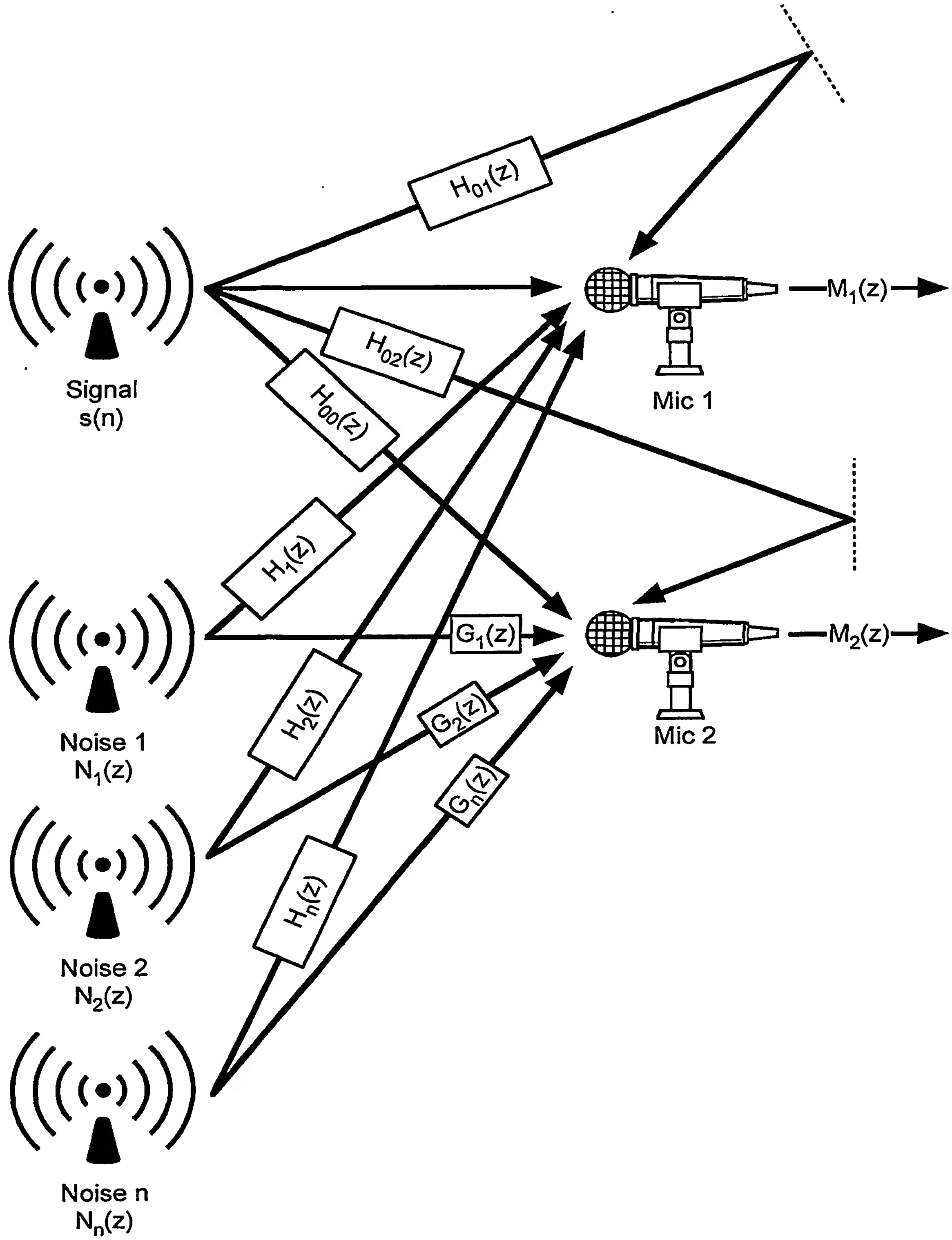

Signal
s(n)
Noise 1
$N_1(z)$
Noise 2
$N_2(z)$
Noise n
$N_n(z)$
$H_0(z)$
$H_1(z)$
$H_2(z)$
$H_n(z)$
$G_1(z)$
$G_2(z)$
$G_n(z)$
Mic 1
Mic 2
$M_1(z)$
$M_2(z)$

Fig. 3



H01(z)
H02(z)
H00(z)
M1(z)
Mic 1
Signal
s(n)
H1(z)
G1(z)
M2(z)
H2(z)
G2(z)
Mic 2
Noise 1
N1(z)
Gn(z)
Hn(z)
Noise 2
N2(z)
Noise n
Nn(z)

Fig. 4

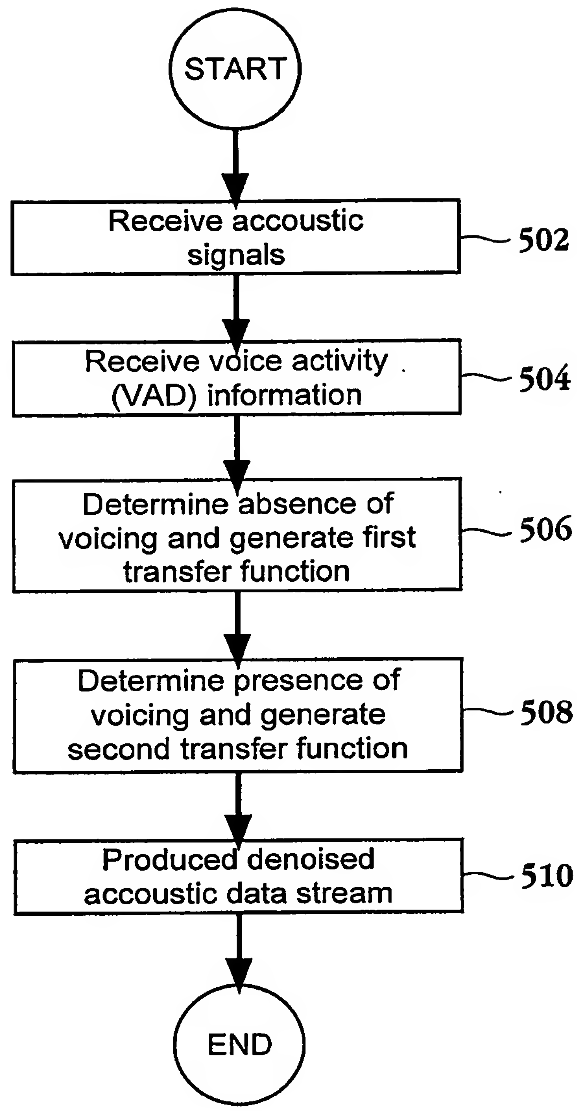

START
Receive accoustic signals
502
Receive voice activity (VAD) information
504
Determine absence of voicing and generate first transfer function
506
Determine presence of voicing and generate second transfer function
508
Produced denoised accoustic data stream
510
END

Fig. 5

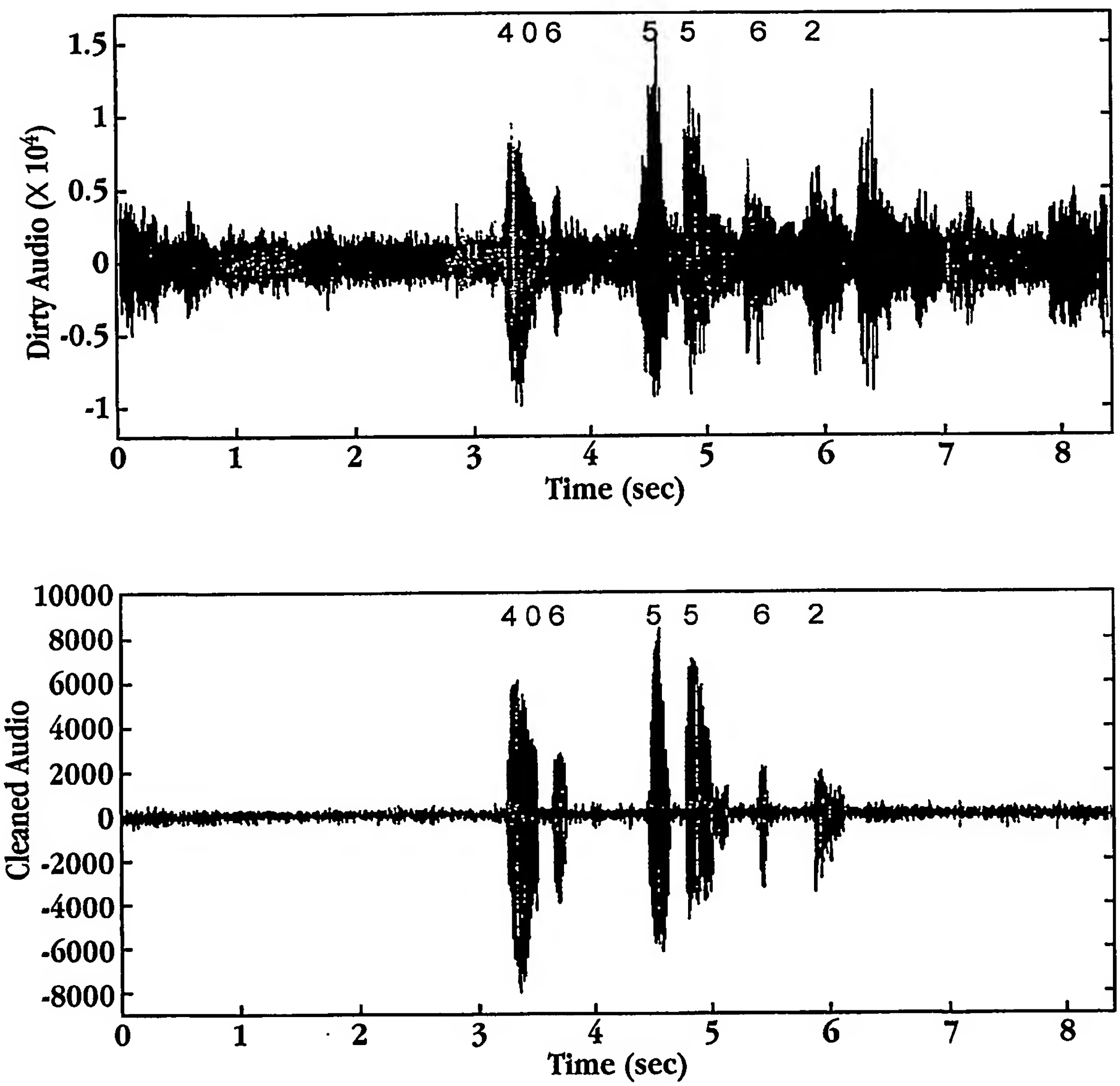

406 5 5 6 2
1.5
1
0.5
0
-0.5
-1
Dirty Audio (X 10⁴)
0 1 2 3 4 5 6 7 8
Time (sec)
406 5 5 6 2
10000
8000
6000
4000
2000
0
-2000
-4000
-6000
-8000
Cleaned Audio
0 1 2 3 4 5 6 7 8
Time (sec)

Fig. 6